

PENGUNAAN *PARTIAL LEAST SQUARE REGRESSION* (PLSR) UNTUK MENGATASI MULTIKOLINEARITAS DALAM ESTIMASI KLOROFIL DAUN TANAMAN PADI DENGAN CITRA HIPERSPEKTRAL

Abdi Sukmono, Sawitri Subiyanto

Program Studi Teknik Geodesi, Fakultas Teknik, Universitas Diponegoro
Jl. Prof. H. Soedarto, SH. Kampus Tembalang, Semarang, email : sukmono35@gmail.com

Abstrak

Klorofil merupakan pigmen yang paling penting dalam proses fotosintesis. Tanaman sehat yang mampu tumbuh maksimum umumnya memiliki jumlah klorofil yang lebih besar daripada tanaman yang tidak sehat. Dalam Estimasi kandungan klorofil tanaman padi dengan airborne hiperspektral dibutuhkan model khusus untuk mendapatkan akurasi yang baik. Citra Hiperspektral mempunyai ratusan band dan julat yang sempit pada setiap bandnya, sehingga mempunyai kemampuan yang cukup baik untuk estimasi klorofil. Akan tetapi karena julat yang cukup sempit ini menyebabkan adanya efek multikolinearitas. Objek dari penelitian ini mengembangkan reflektan in situ menjadi model estimasi kandungan klorofil tanaman padi untuk citra airborne hiperspektral dengan menggunakan metode partial least square regression untuk menghilangkan efek multikolinearitas. Dalam penelitian ini dengan menggunakan teknik hubungan reflektan dan klorofil dipilih band-band yang berhubungan dan efektif untuk estimasi klorofil. Dari hasil seleksi tersebut terpilih 44 band yang efektif untuk estimasi kandungan klorofil daun tanaman padi. Hasil dari penelitian ini menunjukkan metode PLSR dapat menghasilkan model yang cukup baik untuk estimasi kandungan klorofil tanaman padi dengan nilai Koefisien determinasi (R^2) mencapai 0.75 pada PC no 11 dan mempunyai RMSE sebesar 1.44 SPAD unit. Validasi menggunakan data citra airborne hiperspektral menghasilkan RMSE sebesar 1.07 SPAD Unit.

Kata kunci : PLSR, klorofil, hiperspektral, padi

PENDAHULUAN

Latar Belakang

Citra hiperspektral memiliki jumlah kanal yang cukup banyak mencapai ratusan kanal, sehingga mempunyai kemampuan interpretasi yang lebih baik dikanalalkan citra multispektral. Ratusan kanal ini dapat dimanfaatkan sebagai variabel atau parameter estimasi yang cukup baik. Jumlah kanal yang mencapai ratusan kanal dan julat kanal yang sempit ini mampu digunakan untuk estimasi parameter biofisik maupun biokimia seperti klorofil daun. Akan tetapi dengan julat yang sempit (15 nm – 20 nm) menyebabkan adanya efek multikolinearitas yang dapat menjadikan nilai model tidak dapat menjelaskan hubungan antara variabel respon dan variabel prediktor secara baik. Keberadaan multikolinearitas akan menyebabkan varians parameter yang diestimasi akan menjadi lebih besar dari yang seharusnya dengan demikian tingkat akurasi dari estimasi akan menurun.

Menurut Soemartini (2012) penggunaan *Principle Component Analysis* (PCA) dapat mengatasi masalah multikolinearitas. *Principle Component Analysis* (PCA) pada dasarnya bertujuan untuk menyederhanakan variabel yang diamati dengan cara menyusutkan atau mereduksi dimensinya. Hal ini dilakukan dengan cara menghilangkan korelasi diantara variabel bebas melalui transformasi variabel bebas asal ke variabel bebas baru yang tidak berkorelasi sama sekali. Penggunaan metode *Principle Component Regression* (PCR) pernah digunakan oleh Darmawan (2012) untuk estimasi kerapatan daun menggunakan citra hiperspektral melalui data reflektan turunan. Hasil dari pemodelan tersebut dapat menghasilkan nilai model yang cukup baik. Akan tetapi dalam segi komputasi karena mengakomodasi seluruh kanal yang ada (116 kanal) dibutuhkan perangkat komputer yang canggih dalam pemrosesan datanya. Oleh karena itu penggunaan pengembangan metode komponen utama dan reduksi jumlah kanal diperlukan dalam pemodelan estimasi

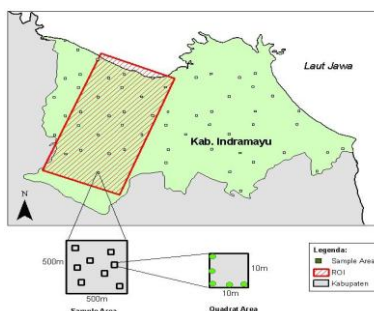
menggunakan data hiperspektral untuk meningkatkan akurasi

Tujuan dari penelitian ini adalah mendapatkan model untuk perhitungan kandungan klorofil tanaman padi pada citra hiperspektral sensor hymap menggunakan pendekatan model Partial Least Square Regression dan mendapatkan panjang gelombang optimal untuk perhitungan Klorofil dengan menggunakan data citra hiperspektral sensor Hymap berbasis data spectrometer in situ

METODOLOGI PENELITIAN

Lokasi Penelitian

Lokasi penelitian dilakukan di sawah beririgasi daerah Indramayu (15 km x 30 km), Jawa Barat (Gambar 1). Pengambilan data dilakukan pada tanggal 25 Juli -13 Agustus 2011. Area ini merupakan daerah lumbung beras dengan beberapa varietas seperti Ciherang, ketan dan IR-42. Selama survei berlangsung, ada 3 fase pertumbuhan padi yang diukur yaitu, fase vegetatif, reproduktif dan pematangan (*ripening*). Data spektral yang diperoleh dikelompokkan sesuai dengan fase pertumbuhannya. Selanjutnya, secara simultan dilakukan pengukuran variabel klorofil (SPAD) dan spectral in situ dengan menggunakan alat FieldSpec 3 untuk spektral in situ dan chlorophyll meter (SPAD-502 Minolta) untuk klorofil.



Gambar 1. Lokasi ROI, *sampling area* (SA), *quadrat area* (QA) dan *hill*

Data Dan Peralatan

- Data

Data yang digunakan dalam penelitian ini adalah:

1. Citra airborne hiperspektral sensor HyMap daerah penelitian yang diambil pada tanggal 13 juli 2011
2. Peta Vektor 1 : 5000
3. Data lapangan yang diambil secara in-situ (Klorofil SPAD, Spektral tanaman, fase tumbuh) di beberapa titik lokasi penelitian pada 25 Juli – 13 Agustus 2011
4. Data posisi sampel area yang diukur dengan menggunakan GPS geodetik.

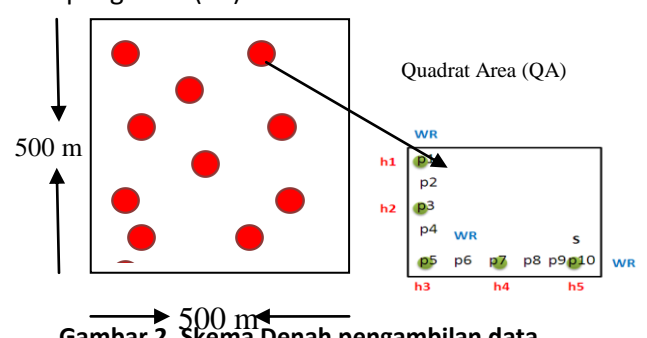
- Peralatan

Peralatan yang digunakan dalam penelitian ini adalah berupa perangkat keras yaitu alat pengukur spectral, GPS geodetic, alat pengukur klorofil dan kamera serta perangkat lunak untuk pengolahan data spasial dan *spectral analysis* menggunakan SAMS 3.2.

Pengolahan Data

Data yang digunakan dalam pekerjaan penelitian ini terdiri dari data Klorofil Lapangan (SPAD), Data kondisi tanaman padi (varietas, foto, ketinggian), Data *Field Spectrometer*, dan data citra *airborne* hiperspektral. Dimana dalam penelitian ini menggunakan *sampling area* (SA) berukuran 500 m x 500 m sebanyak 10 buah, dimana pada setiap *sampling area* terdapat *Quadrat Area* (QA) yang tersebar acak dengan ukuran 10 m x 10 m sebanyak 10 buah/*sampling area*. Pada setiap *sampling area* disetiap pinggir penyikunya terdapat *hill-hill* pengukuran seperti ditunjukkan pada gambar 2 dimana p1 s/d p10 adalah lokasi pengukuran spectrometer, h1 s/d h5 adalah lokasi *hill*, WR adalah lokasi pengukuran *white reference*.

Sampling Area (SA)

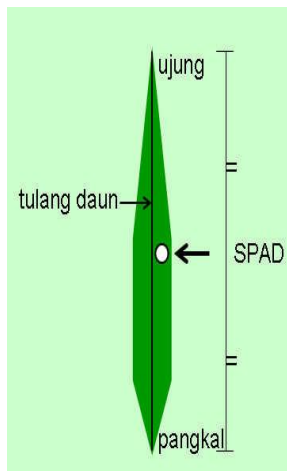


Gambar 2. Skema Denah pengambilan data

Pengambilan Data klorofil

Pengambilan data klorofil dilakuakn mulai tanggal 3 juli 2011- 5 juli 2011 menggunakan alat SPAD. Teknik pengukuran klorofil pada waktu survei lapangan dilakukan sebagai berikut:

- Dalam satu rumpun pengukuran dilakukan tiga kali dan dalam satu kuadrat ada 5 rumpun yang diukur.
- Hitung nilai klorofil rata-rata untuk setiap kuadrat. Adapun pengambilan data klorofil daun dijelaskan pada gambar 3.



Gambar 3. Titik pengukuran klorofil dengan SPAD 502 pada daun padi

Pengambilan Data kondisi tanaman

Pengambilan data kondisi tanaman dilakukan mulai tanggal 30 juni – 5 juli 2011 yang meliputi penentuan growt stage, tinggi tanaman, foto quadrat area.

Pengambilan Data *field spectrometer*

Pengambilan data fieldspectrometer dilakukan tanggal 27 juni 2011-30 juni 2011 menggunakan alat *fieldspectrometer* . Reflekan kanopi daun diukur pada titik pengukuran yaitu rumpun padi dengan field spectrometer pada jam 10.00 – 14.00 dari atas permukaan kanopi daun dengan ketinggian 1, 5 meter dengan sudut kemiringan $23,5^\circ$. Setiap pengukuran dilakukan pengulangan sebanyak minimal lima kali. Reflektan permukaan tanah atau air di sawah pada setiap QA juga diukur reflektannya.

Pengambilan Data Airborne Hiperspektral

Proses pengambilan data airborne hiperspektral dilakukan dengan wahana pesawat terbang yang dilakukan pada tanggal 13 Juli 2011 di daerah sampling area.

Model Partial Least Square Regression (PLSR)

Metode *Partial Least Square Regression* (PLSR) hampir sama dengan metode *Principle Component Regression* (PCR). Jika pada PCR komponen utama dihitung masing-masing pada peubah-peubah X dan peubah Y secara terpisah, maka pada PLSR komponen utama komponen utama disusun berdasarkan konsep NIPALS dengan dekomposisi nilai tunggal (*single value decomposition*).

HASIL DAN PEMBAHASAN

Pengolahan citra

Dalam pemotretan airborne hiperspektral di dapat 4 lajur citra. Setiap lajur citra harus diproses tersendiri dalam proses georeferencing dan koreksi geometrik. Hal ini dilakukan karena setiap lajur mempunyai kesalahan-kesalahan yang berbeda-beda seperti yang ditunjukkan pada gambar 4.



Gambar 4. Lajur Citra Hymap Indramayu Terkoreksi dan Hasil Mosaik Citra

Tabel 1 Koreksi geometrik dari setiap lajur

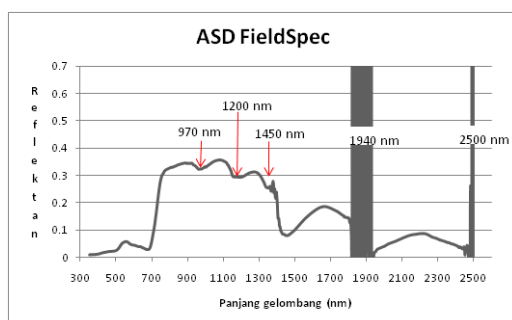
No Lajur	Jumlah GCP	RMSE
1	7	0.686501
2	7	0.753667
3	7	0.513481
4	7	0.706614

Dari keempat lajur pemotretan data yang menggunakan 7 GCP pada setiap lajurnya di dapat RMSE pada setiap lajur kurang dari 1 piksel seperti yang ditunjukkan oleh tabel 1, dengan demikian nilai ini masuk toleransi atau

transformasi koordinat pada setiap lajur citra 7 titik ini dapat diterima dalam koreksi geometrik dan mendekati kebenaran.

Pemodelan Reflektan In Situ

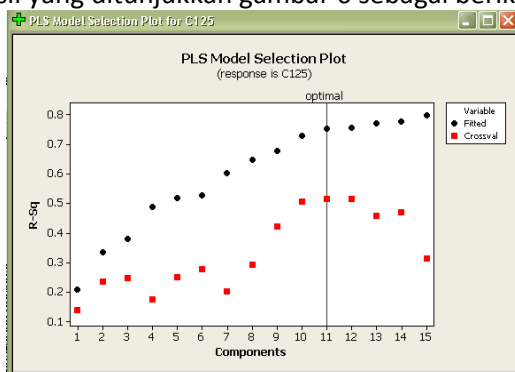
Penggunaan seluruh kanal pada citra HyMap akan memberikan prosesing yang cukup berat pada komputer. Untuk efektivitas estimasi, maka diseleksi kanal-kanal yang mempunyai korelasi terhadap kandungan klorofil dan proses fotosintesis. Gambar 5 menunjukkan pola reflektan vegetasi.



Gambar 5. Pemilihan kanal optimal

Seleksi kanal yang dipilih sebanyak 44 kanal yang meliputi kanal ke 1-24 yang merupakan kanal visible blue (absorpsi klorofil b), hijau (reflektansi klorofil), merah (absorpsi klorofil a), dan Near Infrared (reflektansi vegetasi), kemudian kanal ke 34-36 (water absorpsi vegetasi), kanal ke 54-58 (water absorpsi vegetasi), kanal 66-70 (water absorpsi vegetasi) serta kanal ke 99-105 (reflektansi vegetasi).

Dari ke 44 kanal tersebut dilakuka pemodelan dengan menggunakan Partial least square regression (PLSR) dan validasi silang diperoleh hasil yang ditunjukkan gambar 6 sebagai berikut:



Gambar 6. seleksi model

Dari Gambar 12 terlihat PC no 11 mempunyai nilai terbaik. Hasil dari PC 11 ini dipilih berdasarkan nilai R^2 dan R^2 prediksi tertinggi yang mencerminkan kekuatan model seperti yang ditunjukkan pada gambar 7 berikut.

Component	R-Sq	PRESS	R-Sq (pred)
1	0,208946	461,1	0,141581
2	0,337461	409,164	0,23827
3	0,382132	403,102	0,249555
4	0,489609	441,279	0,178482
5	0,520947	400,91	0,253636
6	0,530472	386,418	0,280615
7	0,60338	427,715	0,203734
8	0,648194	379,584	0,293338
9	0,679008	317,381	0,422886
10	0,73028	264,911	0,506821
11	0,754293	259,385	0,517109
12	0,758625	259,45	0,516988
13	0,774107	290,909	0,458422
14	0,777841	283,214	0,472748
15	0,799515	367,198	0,316397

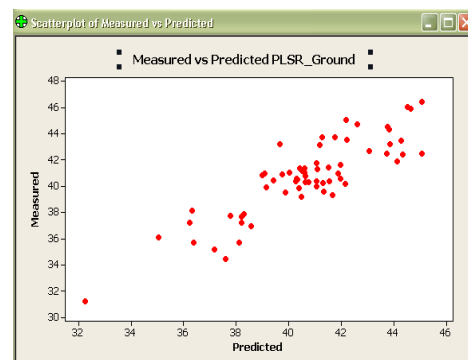
Gambar 7. Hasil regresi PLSR

Selain itu PC no 11 mempunyai nilai PRESS terendah yang menunjukkan kestabilan model jika dimasuki data baru. Hasil model dari PLSR PC no 11 ini mempunyai R^2 sebesar 0.75 dan RMSE sebesar 1.44 unit seperti yang ditunjukkan pada Tabel 3 berikut.

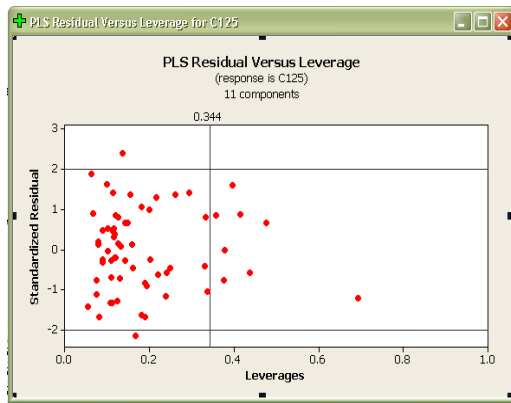
Tabel 3. Hasil PC no 11

PC	R2	RMSE (SPAD Unit)
11	0,754293	1,44

Nilai pendekatan antara nilai prediksi dan pengukuran *in situ* mempunyai pola persebaran yang cukup baik. Perkanalingan nilai predicted dan measures ditunjukkan pada gambar 8.



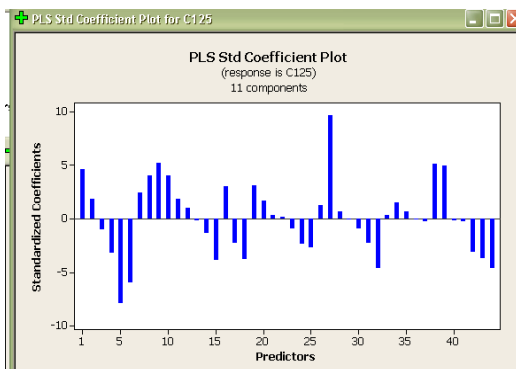
Gambar 8. Hubungan antara klorofil prediksi model dan klorofil pengukuran



Gambar 9. Data pencilan model

Pada gambar 9 menunjukkan jumlah data pencilan dari pengolahan hanya 7 dari total 117 sampel. Hal ini berarti model ini cukup representatif untuk digunakan.

Hasil dari pemodelan PLSR terhadap klorofil menunjukkan bahwa kanal yang optimal untuk estimasi kandungan klorofil daun tanaman padi seperti yang ditunjukkan pada gambar 10 yaitu kanal pada panjang gelombang 455.5 nm, 515.5 nm, 531.2 nm, 981.3 nm, dan 668.5 nm



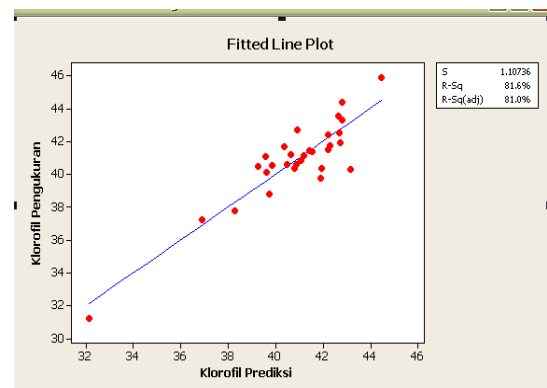
Gambar 10. Perbandingan koefisien plot tiap kanal

Secara ilmiah lima kanal tersebut cukup berhubungan dengan kegiatan klorofil pada tumbuhan. Nilai energi pada panjang gelombang 455.5 nm merupakan panjang gelombang pada posisi biru yang merupakan panjang gelombang yang sensitif dengan klorofil b, sedangkan panjang gelombang 515.5 nm dan 531.2 nm merupakan panjang gelombang hijau yang sangat reflektif terhadap proses fotosintesis, karena pada proses fotosintesis energi pada panjang gelombang hijau akan tidak dibutuhkan. Untuk panjang gelombang 981.3 merupakan posisi pada panjang gelombang yang sensitif terhadap water

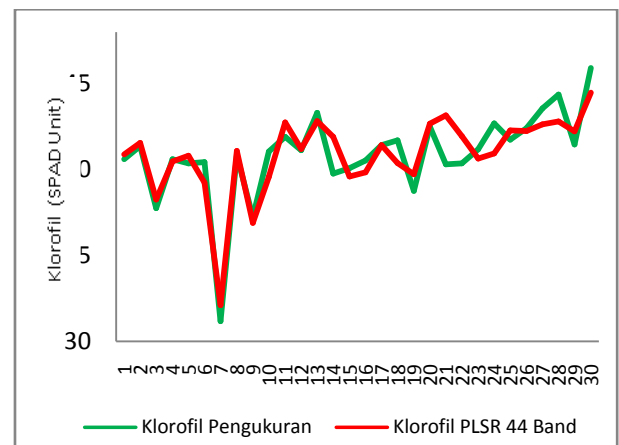
absorpsi pada vegetasi sedangkan panjang gelombang 668.5 nm merupakan puncak absorpsi gelombang merah. Hal ini dikarenakan pada proses fotosintesis dibutuhkan energi pada panjang gelombang biru dan merah. Panjang gelombang merah ini sangat sensitif terhadap klorofil a.

Validasi pada citra airborne Hiperspektral

Validasi model dilakukan menggunakan citra airborne hiperspektral pada 30 sampel titik. Dari 30 sampel tersebut diperoleh RMSE sebesar 1.07 SPAD Unit dan R^2 sebesar 0.816. Hubungan antara klorofil lapangan dan prediksi dapat dilihat pada Gambar. 11 dan Gambar 12.



Gambar 11. Hubungan antara klorofil lapangan dan prediksi

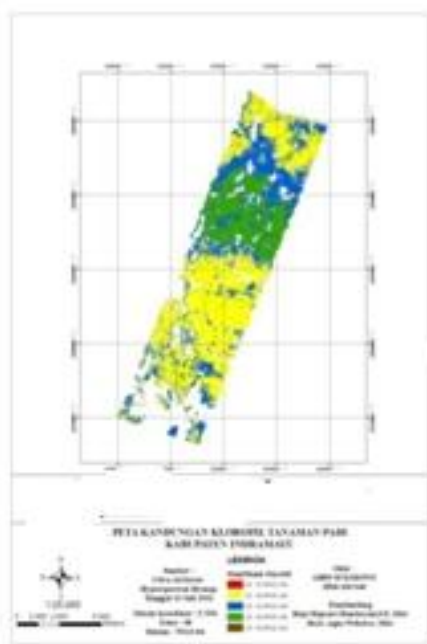


Gambar 12. Grafik perbandingan klorofil lapangan dan klorofil prediksi PLSR 44 kanal

Pada grafik di atas terlihat nilai klorofil prediksi cukup mendekati nilai klorofil sebenarnya. Pola grafik sudah mempunyai pola yang cukup sama.

Peta Distribusi Klorofil

Sasaran akhir dari penyusunan model ini adalah pembuatan peta distribusi (distribution map) klorofil daun pada tanaman padi dari model terpilih. Peta distribusi pada gambar 13 merupakan gambaran secara spasial pada daerah studi. Dalam peta distribusi klorofil ini wilayah terpetakan adalah lahan sawah pada daerah studi. Pembuatan peta distribusi ini sebagai alat untuk pengambilan keputusan terkait pemantauan kandungan klorofil daun tanaman padi.



Gambar.13 Peta distribusi klorofil

PENUTUP

Kesimpulan

Adapun kesimpulan yang dapat di ambil dari penelitian ini adalah sebagai berikut :

1. Metode *Partial Least Square Regression* (PLSR) sangat efektif dalam pemodelan data Hiperspektral karena dapat memanfaatkan banyak kanal dengan menghilangkan efek multikolinearitas, terbukti dengan validasi model menggunakan citra airborne hiperspektral di dapatkan RMSE sebesar 1,07

SPAD unit sehingga model algoritma dapat digunakan.

2. Kanal pada posisi gelombang 455.5 nm , 515.5 nm, 531.2 nm, 981.3 nm, dan 668.5 nm cukup optimal dalam estimasi kandungan klorofil daun tanaman padi.

Saran

1. Diperlukan kajian pemanfaatan metode pendekatan komponen utama lain sebagai pembanding untuk efisiensi pemodelan citra hiperspektral.
2. Dalam pemilihan Band untuk pemodelan dengan metode PLSR dapat dikembangkan penggunaan metode *Genetic Algorithm* (GA) sehingga dapat diperoleh seleksi band yang benar-benar berhubungan terhadap nilai klorofil.

UCAPAN TERIMAKASIH

Terimakasih kepada Pusat Teknologi Inventarisasi Sumber Daya Alam (PTISDA) BPPT atas bantuan data dalam penelitian ini.

DAFTAR PUSTAKA

- Darmawan, Arief.(2012), Pembangunan Model Hiperspektral Untuk Estimasi Produktivitas Vegetasi Padi Berdasarkan Metode Derivatif Regresi Linear, Disertasi Dr., Institut Teknologi Sepuluh Nopember, Surabaya.
- Soemartini.(2012), Aplikasi Principle Component Analysis (PCA) Dalam Mengatasi Multikolinearitas Untuk Mengatasi Investasi di Indonesia Periode 2001.1-2010.4, Jurusan Statistika, Universitas Padjajaran., Jatinangor.
- Subekti, Retno. (2011). Pemanfaatan Software MINITAB Untuk Regresi PLS (Partial Least Square). Jurusan Pendidikan Matematika Universitas Negeri Yogyakarta.
- Tanti, Wiwik Aries. (2013). Perbandingan Metode Kuadrat Terkecil dengan Metode Regresi Komponen Utama Pada Kasus Multikolinearitas. Jurnal Sains Universitas Andalas.